Supplement

Price Index for New One-Family Houses Sold, 1963 to 1967

INTRODUCTION

This report presents the Bureau of the Census new annual index of the price of new one-family houses sold during the years 1963 to 1967. The index is computed on the base 1963=100 and shows an increase of about 10 percent for the 4-year period, with annual gains of 1.1, 2.4, 3.0, and 3.8 points, respectively. (1963 was selected as the base year because this was the first year for which sufficient data were available to compute the index.) The index is computed from data collected by the Bureau of the Census in the Housing Sales Survey, which provides data on the sales prices and selected characteristics of a national sample of new one-family houses built for sale and sold during the years 1963 to 1967.

INDEX OF THE PRICE OF NEW ONE-FAMILY HOUSES SOLD, INCLUDING VALUE OF SITE, 1963 TO 1967

(1963=100)

Year	Index
1963	100.0
1964	101.1
1965	103.5
1966	106.5
1967	110.3

This is an index of the total sales price of new one-family houses built for sale and sold. The sales price includes the value of the site on which the house is built. Since this is an index of sales prices, it reflects not only changes in costs of labor, materials, and land, and selling expense, but also changes in productivity and profit margins in residential construction. The index is designed to measure annual changes in the total sales prices of houses with the same composition of characteristics and not changes in any particular component.

The average of actual sales prices of all new one-family houses sold increased about 24 percent over this 4-year period. Since much of this increase was due to a shift to larger houses with more equipment, it was necessary to adjust for this shift in computing the price index. The device used to adjust for this increase in size and equipment was to identify the characteristics of houses which were considered to be the most important determinants of price differentials and to use the proportions of houses having these characteristics in the 1964-65 period as weights in computing the index. Thus, the index reflects the change over time in the price of houses having a constant "mix" of these characteristics.

The 8 house characteristics used in this analysis are as follows: Size of house (in terms of floor area); number of stories; number of bathrooms; presence or absence of central air conditioning; type of parking facility (garage or no garage); type of foundation (basement or no basement); geographic location (in terms of 12 geographic areas); metropolitan location (inside or outside a standard metropolitan statistical area). There are many other house characteristics for which data are not currently available and for which corresponding adjustments are not now possible. The effects of some of these other characteristics are felt to be minor relative to the effects of the 8 characteristics which are used in computing the index, since the former may be correlated with and, to some extent, accounted for indirectly by the latter. For example, limited tests with data on the incidence of 5 household appliances (stoves, refrigerators, washers, dryers, and dishwashers) in new houses have indicated that, although the effects on the price level of individual houses are sometimes substantial, the effects of the presence of these appliances on the price index are negligible. There are additional characteristics, however, which may have significant effects (such as the principal exterior wall material, whether the house is a detached or row house, and the type of heating system used), and data will be gathered in the future to test these effects.

For purposes of calculating a physical output (or constant dollars) series for residential construction, a price index is needed which excludes site value. On an interim basis, the Bureau of the Census plans to derive a price index for new houses excluding site value, using adjustments drawn from site value data for new houses with mortgages insured by the Federal Housing Administration, which constitute 20 percent of new houses

sold. The interim index is scheduled to be introduced into the constant dollar series for value in place which will be published in Construction Report, Series C30-68-6 "Value of New Construction Put in Place" (copies available from U.S. Government Printing Office, Washington, D.C. 20402, price 10 cents) to be issued in August.

This is the first in a series of price indexes for various categories of construction which the Bureau of the Census plans to develop and publish. The long-range goal of this program is to provide adequate price deflators for all categories of construction for which monthly estimates of value of new construction put in place are published by the Bureau of the Census.

COMPUTATION OF THE CENSUS PRICE INDEX FOR NEW ONE-FAMILY HOUSES

There are two major problems in deriving a price index for new one-family houses: (1) the separation of "pure price" changes from changes in the "quality" of houses and (2) the separation of value of site from the price of the house itself. Clearly, indexes of house prices both including and excluding site value are desirable. The Bureau of the Census is currently publishing only an index including site value, since the necessary data are not yet available to compute an index excluding site value. As noted in the introduction, the feasibility of collecting site value data as part of the Housing Sales Survey is currently being studied.

There are several important characteristics of a house which cause it to be more expensive or less expensive than another house built during the same time period. Since there has been a shift toward larger houses with more equipment during the period under consideration (1963 to 1967), it is necessary to identify these characteristics and estimate their effects on house prices if a measure of "pure price" change during this period is to be developed. The estimating method chosen for this analysis is that of regressing the sales price of the house on its major characteristics to determine the relative importance of these selected characteristics. This method of measuring quality change was first applied to price index analysis by Court in 1938 in a study of automobile prices (reference 7). More recently, this method has been applied to various groups of price data, including automobiles by Griliches (reference 19) and various consumer durables by Gavett (reference 15). The work by Brown (reference 4) on Federal Housing Administration house prices formed the background for the research currently under way at the Bureau of the Census.

There are obviously a great many characteristics of a house which determine its price. From the information available in the Housing Sales Survey, the 8 characteristics listed in table 1 have been selected as those which account for a significant amount of price variability. It should be noted that there are other characteristics for which data are not presently available that may be important price determinants, particularly:

1. Value of site
2. Principal type of exterior wall material
3. Whether the house is a detached or row house (or townhouse)
4. Type of heating system used
5. Presence of one or more fireplaces.

The feasibility of collecting data on these characteristics as part of the Housing Sales Survey is currently being studied.

Of the 8 "characteristics" listed in table 1, there are 2 (geographic location and metropolitan location) which are not directly identifiable as house characteristics but rather are indicative of characteristics which cannot presently be measured directly. Research to date has indicated that the inclusion of the characteristics listed below, in addition to the 8 characteristics listed in table 1, does not seem to improve the regression results significantly:

1. Number of rooms
2. Number of bedrooms
3. Inclusion of any or all of the following appliances in the sale price:
   a. Washer
   b. Dryer
   c. Refrigerator
   d. Stove
   e. Dishwasher.

Available evidence indicates that inclusion of number of rooms and/or number of bedrooms does not improve the results because these characteristics are highly correlated with size of house and number of bathrooms. Likewise, the price variability explained by the 5 appliances listed above seems to be adequately explained by the 8 chosen characteristics.

The 8 characteristics given in table 1 are divided into categories for the purpose of price index analysis. Thus, each house is classified by its location in one of 12 geographic areas, whether or not it has a garage, whether or not it has a basement, etc. Each of these 8 characteristics is treated qualitatively; i.e., each category is treated as a "dummy" independent variable in the regression. Thus, a particular independent variable will have a value of "one" if a particular house falls into its category and "zero" otherwise. One category from each characteristic must be omitted from the regression to avoid an overdetermined system, as noted in table 1.

The form of the regression equation is as follows:

$$Y_i = b_o + b_1 X_{1i} + b_2 X_{2i} + \text{---} + b_{27} X_{27i} + e_i$$

($i = 1, \text{---}, n$), where Y_i is the price of house i;

b_o is the constant term in the regression;

$b_1, b_2, \text{---}, b_{27}$ are the regression coefficients corresponding to the 27 categories included in the regression equation;

X_{ji} equals 1 if house i is in category j and equals O otherwise;

e_i is the "error" or "residual" term in the regression equation;

n is the number of edited observations in the year under consideration. (A separate regression was run for each year, with n varying from about 8900 to 1963 to about 6900 in 1966.)

The regressions for each of the 5 years under consideration produced a coefficient of determination (R^2) of about .70. Thus, this analysis explains about 70 percent of the price variability in the houses included in the Housing Sales Survey.

The constant term in the regression represents the estimated price of a house represented by the omitted categories for each of the 8 characteristics. The coefficients of the independent variables corresponding to those categories which are not represented in the regression constant are indicative of the relative importance of these categories in explaining variability in house prices. If it were possible to completely specify all the statistically independent characteristics of a product which cause it to differ in price from the "basic" version of the product and these could be explained by a strict linear relationship, the regression coefficients would represent estimated unit prices of the characteristics. This is obviously not the case in single-family houses, since there are many interrelated factors which influence the pricing of a house. What has been done in this case is to isolate the available characteristics which are considered to be the most important causes of price differentials and estimate measures of their relative importance.

After the categories of house characteristics have been defined and their regression coefficients have been estimated, index numbers are computed as follows:

$$I_N = \frac{\Sigma P_N Q_{64\text{-}65}}{\Sigma P_{63} Q_{64\text{-}65}},$$

where I_N is the index value for year N,

P_N's are the regression coefficients for the categories in year N,

P_{63}'s are the regression coefficients for the categories in the base year (1963),

$Q_{64\text{-}65}$'s are the proportions of houses represented in the categories in the weighting period (1964-65)

The summations are made over all the categories represented in the regression equation.

This index is the Laspeyres (fixed weight) type with weights as defined above and as shown in table 1. The 1964-65 period was chosen as the weighting period because it was felt that these years were the most representative of residential construction patterns since 1963. 1963 was selected as the base year, since this was the first year for which computation of the index was possible.

The computation procedure outlined above produces the same results as another procedure, which may also serve to illustrate the logic underlying the calculation of the index numbers. The index numbers may be considered as based upon prices of groups of houses, the houses in each group being identical with respect to the 8 characteristics used in the regression analysis. (In theory, there could be 15,552 such groups. Actually, the number of groups is much smaller, because houses with many of the possible combinations of characteristics are not found in practice. For example, there were about 1,400 of these groups represented in the 1963 data.) An estimated price for each group of houses with identical characteristics is obtained from the regression equation (this is not necessarily identical to the actual average price of the houses in the group). These estimated prices are then combined into index numbers, using as weights the proportions of houses in the groups.

Table 1. HOUSING CHARACTERISTICS AND CATEGORIES USED IN CONSTRUCTION PRICE ANALYSIS

(Definitions of the characteristics are given in supplement A. The States included in each of the 12 geographic areas are listed in supplement B.)

Characteristics and categories	1964-65 proportions (percent)
Size of house (floor area)	
Area 1 (less than 1,000 square feet)..	10.2
Area 2 (1,000 to 1,199 square feet)[1]..	20.9
Area 3 (1,200 to 1,399 square feet)...	18.4
Area 4 (1,400 to 1,599 square feet)...	13.8
Area 5 (1,600 to 1,799 square feet)...	11.6
Area 6 (1,800 to 1,999 square feet)...	9.1
Area 7 (2,000 to 2,199 square feet)...	6.8
Area 8 (2,200 to 2,399 square feet)...	4.2
Area 9 (2,400 square feet or more)....	5.0
Number of stories	
1-story[1].....	73.0
Split-level.....	11.4
2 stories or more.....	15.6
Number of bathrooms	
Less than 1 1/2 bathrooms.....	23.2
1 1/2 or 2 bathrooms[1].....	61.0
2 bathrooms or more.....	15.8
Central air conditioning	
Present.....	22.5
Not present[1].....	77.5
5. Type of parking facility	
Garage[1].....	68.6
No garage.....	31.4
6. Type of foundation	
Basement.....	40.9
No basement[1].....	59.1
7. Geographic location	
New England.....	6.0
Middle Atlantic[1].....	7.5
East North Central.....	16.6
West North Central.....	8.2
South Atlantic (except Florida).....	16.2
Florida.....	7.5
East South Central.....	2.8
West South Central.....	10.6
Mountain (except Arizona and Nevada)...	3.0
Arizona and Nevada.....	2.7
Pacific (except California and Hawaii).	2.5
California and Hawaii.....	16.4
8. Metropolitan location	
Inside an SMSA[1].....	75.4
Outside an SMSA.....	24.6

[1]Variable omitted from the regression equations.

Supplement A. Variables Used For Construction Price Analysis

Dependent Variable: The sales price of the house (including value of site) is the price agreed upon between the buyer and seller at the time the first sales contract is signed or deposit is made for the house. The houses included in the regressions for a particular time period are those reported as sold during that period (whether or not the house has been started). Only houses built for sale are included; i.e., houses built for rent or for the exclusive use of the owner are excluded. (See reference 28.)

Independent Variables:

1. Size of house (floor area).

Floor area (based on exterior dimensions) includes all areas with finished walls, floors, and ceilings, including such finished areas in basements and attics. Where floor area for a house is reported in interior dimensions, the figure is converted to exterior dimensions by multiplying by a standard conversion factor of 1.08. For purposes of this analysis, a standard factor of 1.04 is used to convert figures to exterior dimensions where it is not known whether the reported area is based on exterior or interior dimensions. After conversion to exterior dimensions, the reported areas are grouped into nine categories with "less than 1,000 square feet" as the lowest, "2,400 square feet or more" as the highest, and seven intermediate categories with 200 square-foot intervals.

Number of stories.

For the purpose of this analysis, houses are divided into categories of 1-story, 2 stories or more, and split-level. The 1-story category includes a small number of houses which are reported as having 1 1/2 stories. A split-level house is generally considered as one having floors on more than one level when the difference in some floor levels is less than one story. Distinctions between half stories and split-levels may vary slightly from region to region.

3. Number of bathrooms.

A full bathroom is defined to include a toilet, a basin, and a bathtub and/or shower, while a half-bathroom is defined to include a toilet and a basin. For purposes of this analysis, houses are grouped into three categories: less than 1 1/2 bathrooms, 1 1/2 or 2 bathrooms, 2 bathrooms or more.

4. Central air conditioning.

This refers to whether the house did or did not have central air conditioning at the time it was built.

5. Type of parking facility.

For the purpose of this analysis, houses are divided into two categories: those having a garage (for one or more cars), those having no garage (including houses with carports).

6. Type of foundation.

For the purpose of this analysis, houses are divided into two categories: those having a full or partial basement, those built on a slab or having only crawl space.

7. Geographic location.

The Housing Sales Survey covers sales of houses in a national sample of land areas. The geographic areas used for the regressions are given in supplement B. These particular areas were chosen because it was felt that this was the finest breakdown possible for isolating geographic price differentials.

8. Metropolitan location.

Whether or not the house is in a standard metropolitan statistical area is determined by the 1960 definition of SMSA's (see reference 27).

Weights: The regressions in this analysis are weighted regressions, using weights based on the sampling rates used in the Housing Sales Survey.

Editing: The houses built for sale which are included in the Housing Sales Survey are edited for this analysis using a two-stage procedure. First, those houses for which the sales price or any of the eight characteristics are not reported are rejected. Second, the reported figures on sales price and floor area are checked for extremeness. Houses falling outside the following three sets of limits are excluded:

Price per house	Price per square foot of floor area	
	Lower limit	Upper limit
Less than $15,000.	$5.00	$18.00
$15,000 to $19,999	8.00	20.00
$20,000 or more...	9.00	25.00